(4)

ARL-SYS-TM-109

AR-005-522

AD-A221 280

DEPARTMENT OF DEFENCE

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION

AERONAUTICAL RESEARCH LABORATORY

MELBOURNE, VICTORIA

Aircraft Systems Technical Memorandum 109

IMAGE SEGMENTATION FOR TARGET AIRCRAFT DETECTION (U)

by

J. Sandor

and

R.C. Creek

Approved for public release.

AUGUST 1988

AR-005-522

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
AERONAUTICAL RESEARCH LABORATORY

Aircraft Systems Technical Memorandum 109

IMAGE SEGMENTATION FOR TARGET AIRCRAFT DETECTION

by

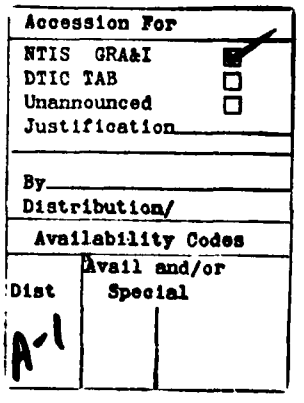
J. Sandor and R.C. Creek

SUMMARY

A two stage method is presented for segmenting digitized images into aircraft and background. The approximate aircraft position is determined from the edge distribution and then each pixel within this region is classified by statistical comparison with the closest local region outside the search area.

POSTAL ADDRESS: Director, Aeronautical Research Laboratory,
P.O. Box 4331, Melbourne, Victoria, 3001, Australia

CONTENTS



1. Introduction

The automatic discrimination of target objects from the background in greyscale images is an important element in a large number of applications ranging from target designators through medical imaging to robot vision systems. In many typical applications the objects of interest in an image have distinguishing characteristics which are readily definable in terms of grey level contrast with respect to the background. In these instances histogram analysis and thresholding constitute the appropriate technique for producing the binary partitioning of the image. It is particularly well suited for those applications where the greyscale distribution of the image is bimodal [1] and the background contributes to only one of the modes. In other approaches the image gradient provides the mechanism for segmentation whilst in yet others it is the statistical characteristics of image texture. The former is applicable to recognition tasks in such areas as robot vision where controlled illumination ensures sharp contrast boundaries [2] whilst the latter is well suited for segmentation of images with constituent objects of varying granularity [3]. An excellent survey of the current state-of-the-art in target recognition is presented in reference [4].

In our application we are interested in assessing the accuracy of manual tracking of aircraft by analysis of digital images of the operator's field of view. This requires the extraction of target pixels from the image, calculation of the spatial distribution of the target (up to second moments) and estimation of the instantaneous aim point accuracy from the relative position of the cross-hairs and the target. The aircraft of likely interest range over a wide spectrum of types, at various scales and orientations and against the full gamut of daytime sky conditions and illuminations as typified by the example in Figure 1. Under these constraints the conventional methods of image segmentation mentioned above are not applicable as the target may have positive or

negative contrast, internal shadows and highlights leading to a unimodal greyscale histogram. Moreover the presence of internal edges as well as the discontinuous nature of the edges makes it difficult to devise a robust method for discriminating the target silhouette, whilst the number of possible aircraft types/orientations make approaches based on prior training unattractive.

To accommodate the constraints of the present application we have implemented a composite segmentation technique based on edge detection and statistical pixel classification. Under the assumption that man-made objects have in general higher image gradients than natural features, thresholding of the image gradient leads to the delineation of a region of the image within which the target is likely to be found. Final segmentation involves approximating the latter region by the rectangle inscribed by the second moment ellipse of the edge distribution and comparing the grey level of each pixel within this region to the average grey level of the "closest" neighborhood outside the region. If the difference between the pixel and the average neighborhood grey level is significant then the pixel is classified as belonging to the target. The technique permits variation within both target and background pixels and only requires that local inter-class differences be significant for successful segmentation.

The paper is organised as follows: the data acquisition system and image preprocessing are described in Section 2 whilst edge detection and target region designation are described in Section 3. The essential features of pixel classification together with examples of segmented images are presented in Section 4. Quantitative assessment of the effectiveness of the approach as well as its limitations are discussed in Section 5.

2. Image Acquisition and Preprocessing

The images of interest are acquired by real time digitization of the output of a CCIR standard video camera monitoring the field of view of an optical tracker. The acquisition system has variable resolution such that at the highest resolution a 64×64 array corresponding to one eighth of the camera field is captured whilst at the coarsest resolution the 64×64 data array spans the full field. The choice of resolution is under computer control and is adjusted on the basis of range sensor information such that the target aircraft image, over the range of relevant aircraft types, will not span more than 20% of the captured image. Image greyscale is quantized to 64 levels.

The image contains not only the target and the background but also the cross-hairs used by the operator to boresight the tracker. The location and extent of the cross-hairs have to be established such that the aim point can be estimated and the cross-hair lines eliminated to avoid confounding the edge detection phase of the segmentation process. The motion of the cross-hair with respect to the camera field of view (caused by minor optical misalignment) necessitates the estimation of the cross-hairs at each instant an image is captured.

For the purposes of exposition let the grey level at (i, j) in the image array be represented by $I(i, j)$ and let the coordinates of the camera pixels be denoted by (u, v). Then

$$i = \lfloor \frac{u - u_0}{m} \rfloor; \qquad u = u_0, u_0 + m, \cdots, u_0 + 63m$$
$$j = \lfloor \frac{v - v_0}{m} \rfloor; \qquad v = v_0, v_0 + m, \cdots, v_0 + 63m$$

where (u_0, v_0) is the bottom left hand corner of the data window, m is an integer power of 2 defining the capture resolution and $\lfloor a \rfloor$ is greatest integer less than or equal to a.

Crosshair estimation then proceeds from the prior knowledge that the four lines are orthogonal and close to horizontal/vertical. The positions of the four line segments are first determined and the cross-hair centre found by interpolation. The position of the left-hand near horizontal line segment is found, as shown in Figure 2, by examining the image region near the left border of the image. Under the assumption that this line segment is approximately horizontal it is highly probable that it will lie along a single row over this image region. Hence summing the grey levels of all the rows in this region will reduce the influence of noise and increase the effect of this cross-hair segment. As the cross-hair is always darker than the surrounds the row along which it lies can be distinguished by convolving the row sums(S_i) with the mask 1,-2,1 such that the resultant row sum value $S_n := S_{n-1} - 2S_n + S_{n+1}$. The position of the maximum S_n value defines the position of the cross-hair segment. Similarly the positions of the other three cross-hair segments are determined.

3. Edge Detection and Target Region Designation

The first stage of the segmentation process involves a low level partition of the image into two regions one of which is designated as having a high likelihood of containing the target. Under the assumption that man-made objects account for the highest contrast within the image, the image gradient is the appropriate quantity of interest. Choosing an appropriate threshold, the image gradient is sliced such that the region enclosing the high contrast edges can be determined.

The gradient $G(i,j)$ is computed via the Roberts cross gradient operator which, although not as robust against noise as the Sobel operator[5] is computationally less burdensome and was found experimentally to perform equally well in our application. The threshold at which the gradient is sliced is calculated as the 97.5% point of the distribution of G. This is justified on the basis that images are generally scaled so that the target spans approximately 15% of them and the targets them-

selves are typically observed to have an aspect ratio of 5:1 which together infer the target will have a perimeter spanning about 4% of the image pixels. Since the target contributes the strongest edges (including internal edges due to shadows) it is reasonable to assume that over the range of aircraft sizes, the upper 2.5% of the gradient distribution is contributed principally by the target. Slicing a large number of images over a range of aircraft types/sizes confirmed the suitability of the threshold level. Examples of target edge maps obtained by the process are depicted in Figure 3. It will be noted that the edges are not connected so that derivation of a target silhouette (or equivalently a binary template), which is the ultimate quantity of interest, will in general require excessively complex reasoning.

The target edges contain significant information in that they delineate the region of the image containing the target. In the present case the region is specified by the second moment ellipse of the distribution of the thinned binary edge map. Thinning is necessary to minimise the bias that arises from edge broadening due to the approximate nature of the gradient operator and is accomplished by an implementation of the method described in [6].

Denoting the (thinned) binary edge map by $G_T(i,j)$, the parameters of the second central moment ellipse about the centroid $(\bar{i},\bar{j})$ of the edges is given by [7]

$$\text{semi} - \text{major axis} = a$$
$$= \left(\frac{m_{20}+m_{02}+[(m_{20}-m_{02})^2+4m_{11}^2]^{1/2}}{m_{00}/2}\right)^{1/2}$$
$$\text{semi} - \text{minor axis} = b$$
$$= \left(\frac{m_{20}+m_{02}-[(m_{20}-m_{02})^2+4m_{11}^2]^{1/2}}{m_{00}/2}\right)^{1/2}$$

where

$$m_{00} = \sum\sum G_T(i,j)$$
$$m_{10} = \sum\sum iG_T(i,j)$$
$$m_{01} = \sum\sum jG_T(i,j)$$
$$m_{11} = \sum\sum i'j'G_T(i,j)$$
$$m_{20} = \sum\sum i'^2G_T(i,j)$$
$$m_{02} = \sum\sum j'^2G_T(i,j)$$
$$\bar{i} = m_{10}/m_{00}, \quad \bar{j} = m_{01}/m_{00}$$
$$i' = i - \bar{i}, \quad j' = j - \bar{j}$$

with the sums taken over all image points.

If one considers a coordinate frame $O'X'Y'$ coincident with the principal axes of the ellipse then a^2, b^2 are the variances of the spread of non-zero elements of G_T in the X', Y' directions respectively. To account for aircraft attitudes that lead to non-convex silhouettes an ellipse with semi-major, -minor axes of length $2.5a$, $2.5b$ is chosen as the region within which the aircraft is likely to be found. Since we are dealing with an ensemble of images then by the central limit theorem the edge pixels of G_T are distributed with a bivariate normal density $\frac{1}{2\pi ab}\exp\frac{1}{2}(\frac{x^2}{a^2} + \frac{y^2}{b^2})$ and so it follows that $w = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ has a χ_2^2 distribution. In such a case the latter ellipse defines the 95% probability region.

4. Pixel Classification

The remaining element of image segmentation is the classification of the image pixels within the ellipse defined above. To accomplish this we first construct a rectangle circumscribing the ellipse and then consider the set of pixels parallel to

but outside the rectangle as shown in Figure 4. Considering these pixels as the set $\{a_k\}$, we assign to each element the mean $(\hat{\mu})$ and variance $(\hat{\sigma}^2)$ of the image grey level of its 5×5 neighborhood. This neighborhood of a point a_k is displayed in Figure 5 within the context of the neighboring members of $\{a_k\}$ which are shaded. Then to classify each pixel within the ellipse with respect to the sky or background, its grey level is compared to the distribution of the region about the closest a_k. If

$$|I(i,j) - \hat{\mu}_k| \geq 2.25\hat{\sigma}_k$$

then the pixel at (i,j) is classified as belonging to the aircraft. Repeating this for each pixel within the ellipse leads to a binary partition of the original image.

One can see that the classification procedure permits variation within the target and the background, requiring only that "local" differences between the two classes be present. Unlike the histogram approach account is taken of spatial variation. This is well illustrated by considering the image in Figure 3(a). The grey level histograms of the image, detected aircraft and the background are shown in Figure 6. Clearly the target aircraft is masked by the background in the image histogram, disallowing the use of a histogram segmentation process. By contrast the pixel classification technique readily extracts the aircraft as shown by the segmentation of Figure 7. Similarly, Figure 8 depicts the captured image, the grey level histogram, the sliced gradient and the segmented binary image of the actual aircraft.

5. Performance

In assessing the performance of the method presented we need to quantify how well the extracted binary template matches the actual aircraft shape within the image. A common method for examining shape fidelity of two templates is that of

cross-correlation analysis [1]. If $B(i,j)$ is the $N \times N$ array corresponding to the template extracted by pixel classification and $A(i,j)$ is the $N \times N$ array representing the reference aircraft shape then the normalized cross-correlation between the two is defined as

$$R(i,j) = \frac{\sum_{l=1}^{M}\sum_{m=1}^{M} B(l,m)A(l+i,m+j)}{\left[\sum_{l=1}^{M}\sum_{m=1}^{M} B^2(l,m)\right]^{1/2}\left[\sum_{l=1}^{M}\sum_{m=1}^{M} A^2(l,m)\right]^{1/2}}$$

for $1 \leq i,j \leq N - M$, where M is the size of the smallest array that will support all the non-zero elements of A and B.

The above cross-correlation function was calculated for five representative images of actual aircraft processed by the method of Section 4. In each case the binary reference shape was extracted manually. These are depicted in Figure 9 together with the corresponding detected binary templates. Figure 9(a) corresponds to the image in Figure 8 and exhibits a maximum cross-correlation (equal to 0.7) when there is zero spatial lag. Similarly the other cases showed maximum cross-correlation at zero lag with maximum values ranging from 0.93 (for pair in Figure 9(d)) to 0.74 (Figure 9(e)). Moreover cross-correlation of templates from different pairs showed maximum values of less than 0.4 indicating that even though the mean and variance of the cross-correlation coefficient is not known for each pair of template and shape, $R(i,j)$ is a reasonable metric for shape fidelity.

The principal limitation of the procedure described in Section 3 and 4 is that it will not handle situations where the object of interest is of such aspect and extent that it extends beyond the ellipse enclosing the edge pixels. In this case the pixel classification procedure would assume that the extreme aircraft sections protruding beyond the ellipse would constitute the local background and in comparing these

regions with aircraft regions inside the ellipse would note their similarity and so incorrectly classify these enclosed pixels as belonging to the background. Similarly if two disjoint objects are present in the image then the procedure will, in general, only extract the one with the strongest edge contrast. Such images can occur for low flying aircraft when they are viewed against terrain or for aircraft flying among high contrast cloud formations. Consideration of the pixel classification process indicates that if the variation, across the search ellipse, of the background intensity (due to cloud) is greater than the spread of the sky grey levels this may produce distracting noise pixels in the classified binary image. The effect of this noise will begin to deleteriously affect the segmentation when the intensity variation approaches twice the spread in the sky grey levels. Our experimental work indicates that such high contrast cloud backgrounds mainly occur close to the sun and so are rarely encountered in this application.

Finally the extent of any ellipse approximating the extracted template will be influenced by the presence of outliers. However their effect can be minimised ignoring the extreme outliers beyond the ellipse boundary and then re-calculating the ellipse. This process is repeated until the size of the ellipse stabilizes.

6. Concluding Remarks

A two pass procedure has been presented for extraction of aircraft from images. It relies on the principle that man-made objects will contribute the greatest edge contrast and that they differ significantly in intensity from the background (i.e. they can be visually discerned). Instead of requiring that the sliced image gradient have a connected set of edge pixels defining the aircraft silhouette, the sliced image gradient is used to define a reduced search region (an ellipse) which is then examined to detect pixels which differ significantly from the adjacent backgound regions outside the ellipse. In this way the processing can accommodate grey level variation within the

aircraft and only requires that it be significantly different from the local background. Thus unlike conventional histogram analysis spatial distribution effects are included and furthermore the process is not unduly sensitive to selection of the gradient threshold because it is two pass.

Experimental assessment against actual aircraft models verified the validity of the procedure.

References

[1] **Rosenfeld, A and Kak, A C**, *Digital Picture Processing*, Academic: New York, USA (1976)

[2] **Gonzales, R C and Safabakhsh, R**, 'Computer vision techniques for industrial applications and robot control,' *IEEE Computer*, Vol. 15, No 12 (January 1982)pp 17-32

[3] **Haralick, R M**, 'Statistical and structural approaches to texture,' *Proc. IEEE*, Vol. 67, No. 5 (May 1979) pp 786-804

[4] **Bhanu, B**, 'Automatic target recognition: state of the art survey,' *IEEE Trans Aerospace and Electronic Systems*, Vol. AES-22, No. 4 (July 1986) pp 364-79

[5] **Abdou, I E and Pratt, W K**, 'Quantitative design and evaluation of enhancement/thresholding edge detectors,' *Proc. IEEE*, Vol. 67, No. 5 (May 1979) pp 753-763

[6] **Pratt, W K and Kabir, I**, 'Morphological binary image processing with a local neighborhood pipeline processor,' *Frontiers in Computer Graphics*, Ed. T.L. Kunii, Springer-Verlag, Tokyo,Japan(1985)

[7] **Teague, M R**, 'Image analysis via the general theory of moments,' *J. Opt. Soc. Am.*, Vol. 70, No. 8, (August 1980) pp 920-930

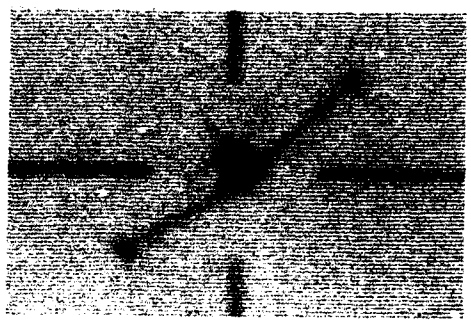


FIGURE 1. TYPICAL IMAGE OF AN APPROACHING AIRCRAFT

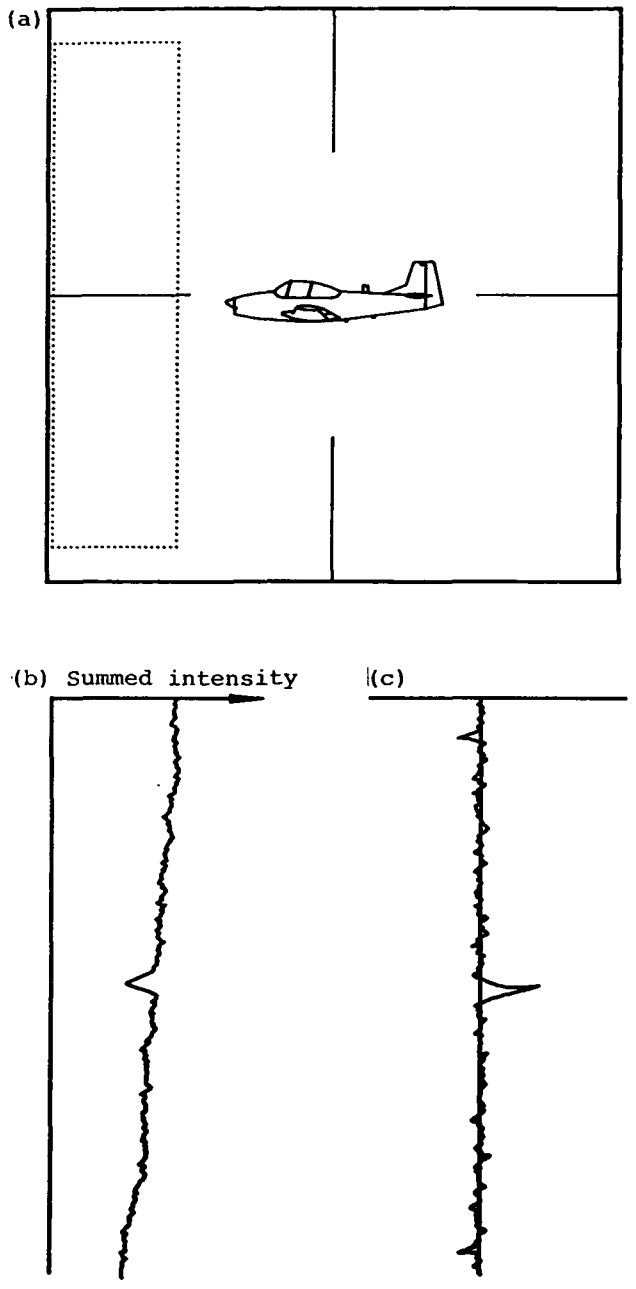


FIGURE 2. DIAGRAMMATIC REPRESENTATION OF THE TECHNIQUE USED TO DETERMINE THE POSITION OF THE LEFT-HAND, HORIZONTAL GRATICULE SEGMENT. ROW SUMMATION OF GREY LEVELS IN THE REGION ENCLOSED BY THE DOTTED RECTANGLE IN (a) YEILDS THE INTENSITY PROFILE IN (b). CONVOLVING THIS INTENSITY PROFILE WITH THE MASK 1,-2,1 YIELDS (c) WHERE THE MAIN PEAK GIVES THE GRATICULE SEGMENT POSITION.

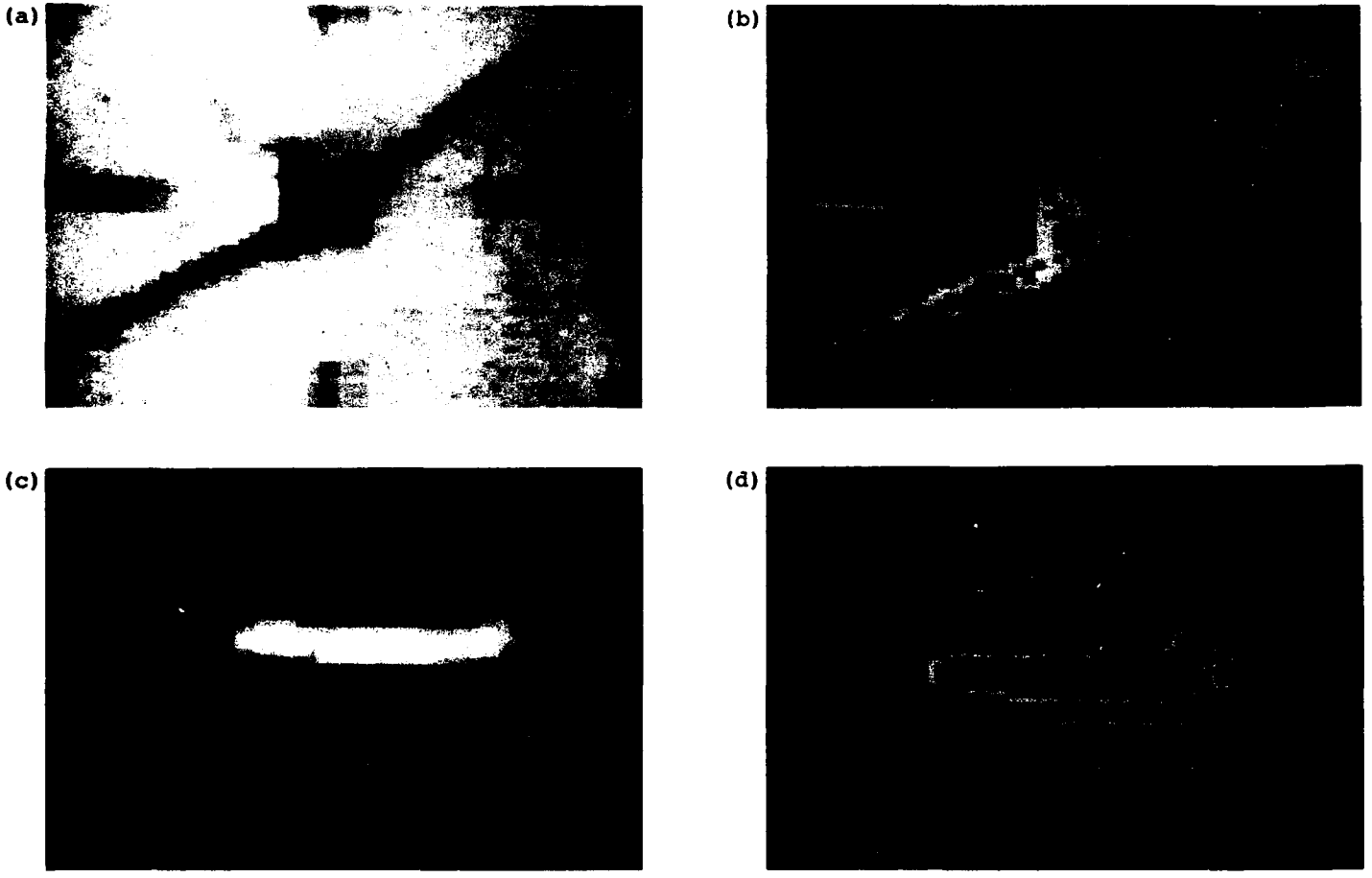


FIGURE 3. EDGE MAPS CALCULATED BY APPLYING THE SOBEL OPERATOR TO THE CAPTURED DIGITIZED IMAGES (a) AND (c) SHOWN IN (b) AND (d) RESPECTIVELY.

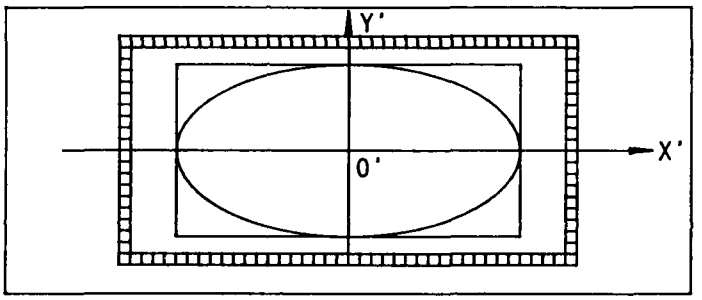


FIGURE 4. THE SET OF LOCAL-BACKGROUND PIXELS $\{a_k\}$ WHICH ARE USED TO SEGMENT THE ENCLOSED PIXELS INTO AIRCRAFT AND BACKGROUND REGIONS.

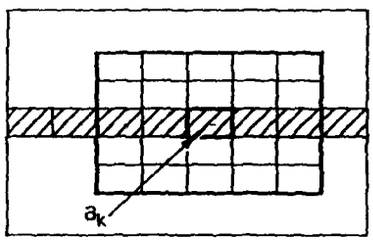


FIGURE 5. THE NEIGHBORING 5x5 REGION OF THE POINT a_k USED TO DETERMINE THE LOCAL INTENSITY VARIATION OF THE BACKGROUND.

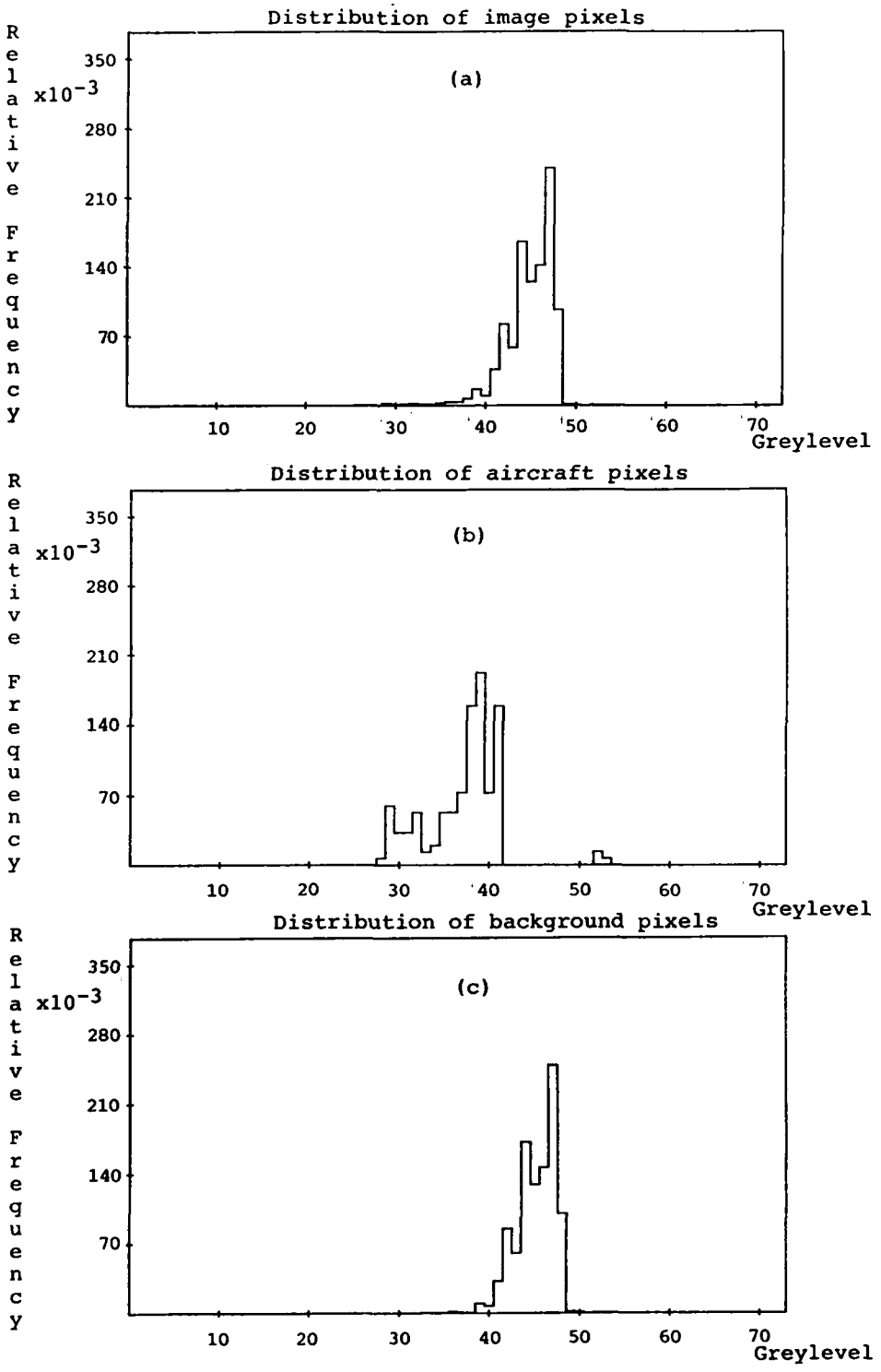


FIGURE 6. (a) THE GREY LEVEL HISTOGRAM OF THE IMAGE SHOWN IN FIGURE 3a AND THE SEPARATE HISTOGRAMS OF THE AIRCRAFT (b) AND THE BACKGROUND (c).

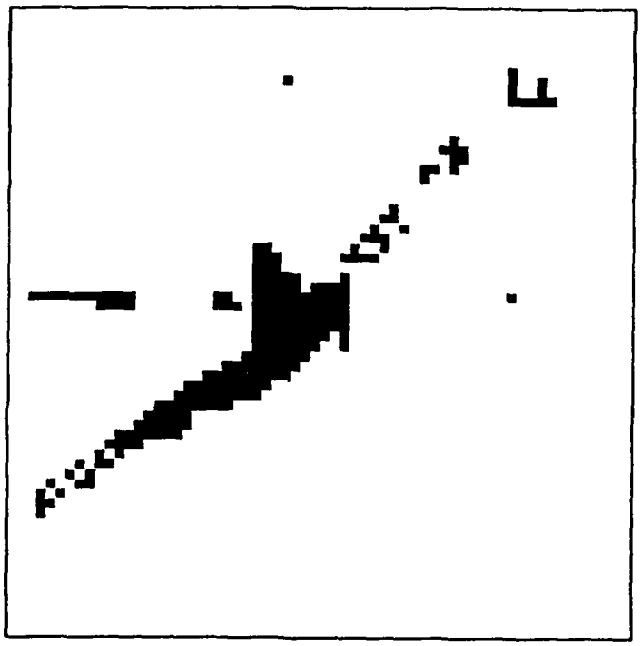


FIGURE 7. SEGMENTED BINARY IMAGE OF THE GREY-SCALE IMAGE SHOWN IN FIGURE 3a. THE BLACK PIXELS DEFINE THE AIRCRAFT POSITION.

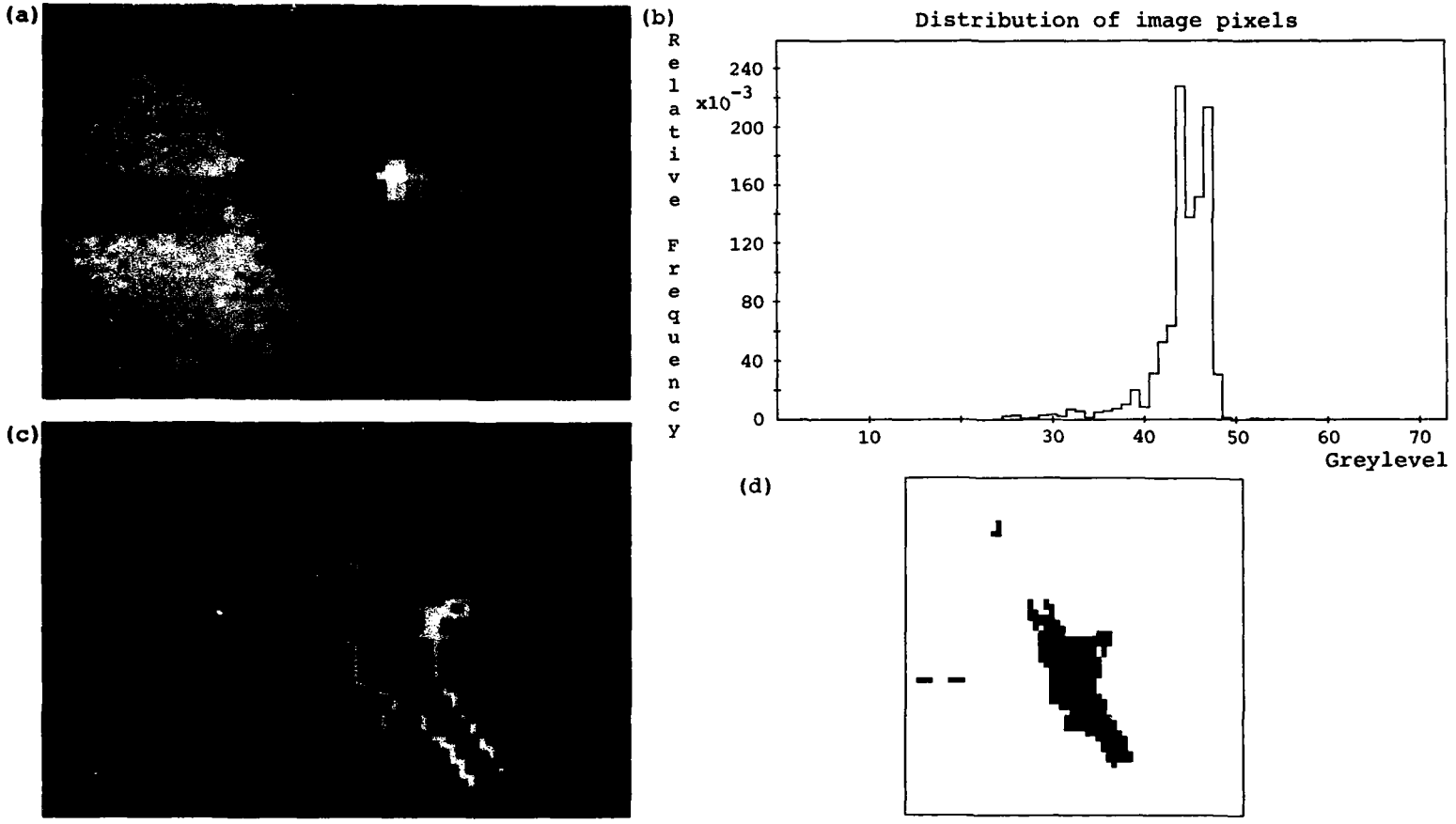
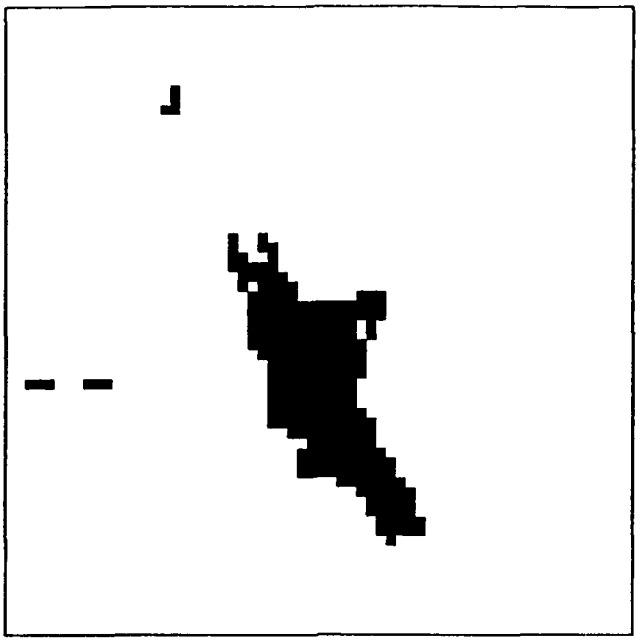


FIGURE 8. PROCESSING OF A TYPICAL CAPTURED IMAGE (a), SHOWN WITH ITS GREY-LEVEL HISTORGRAM (b). FIRSTLY THE SLICED AND THINNED IMAGE (c) IS DETERMINED AND IS USED TO DEFINE THE REGION TESTED BY THE PIXEL CLASSIFICATION PROCESS TO PRODUCE THE SEGMENTED BINARY IMAGE (d).



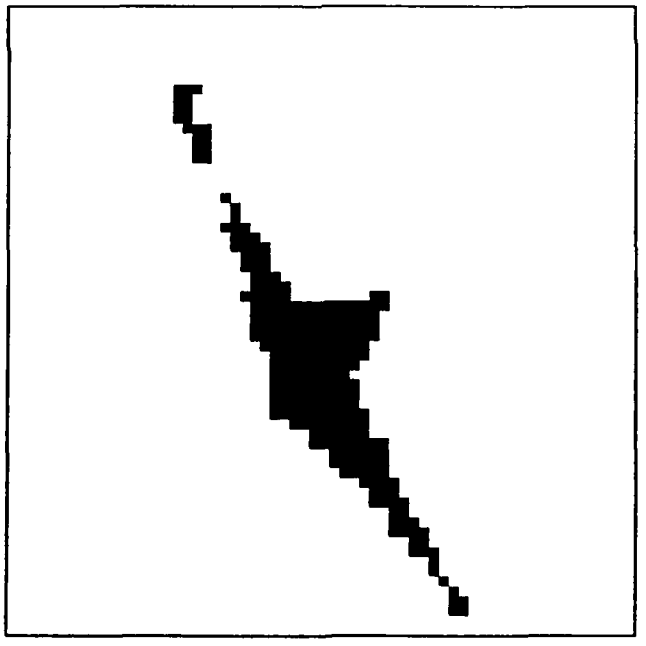
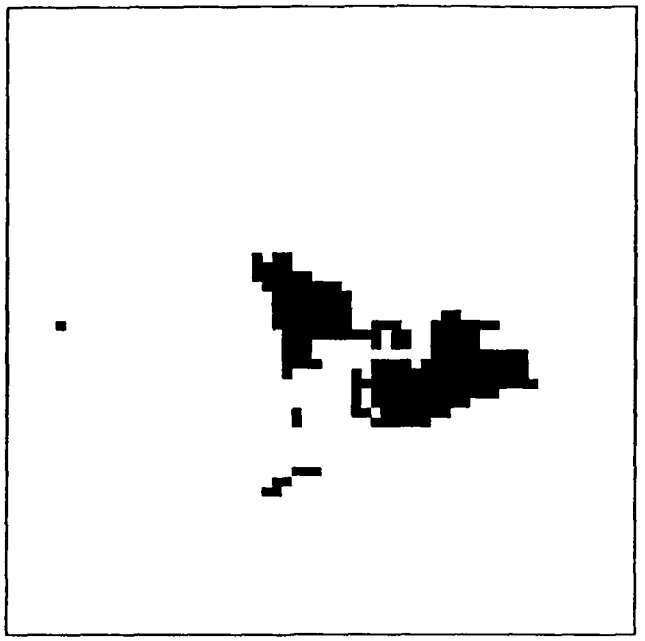


FIGURE 9a. PERFORMANCE OF THE PROCESSING SHOWN FOR FIVE CAPTURED IMAGES WHERE THE TOP ELEMENT IN EACH OF THE DISPLAYED IMAGE PAIRS WAS EXTRACTED USING THE PIXEL CLASSIFICATION PROCESS AND THE LOWER ELEMENT WAS EXTRACTED MANUALLY.



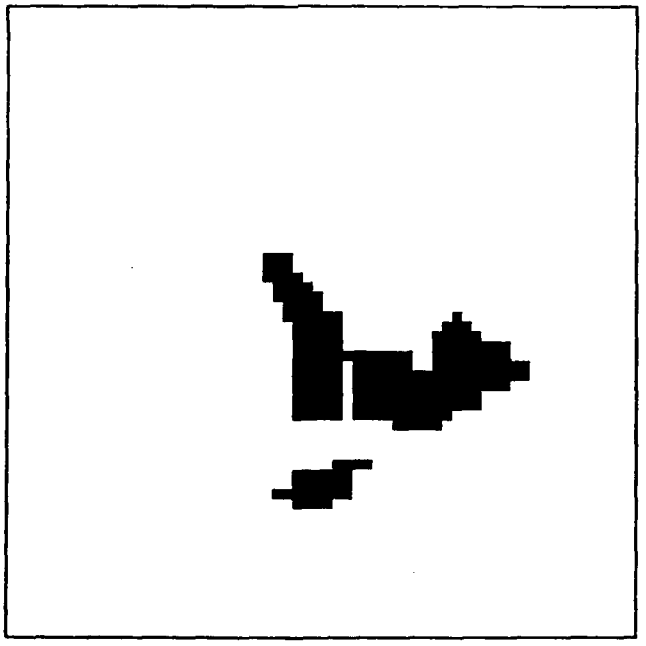
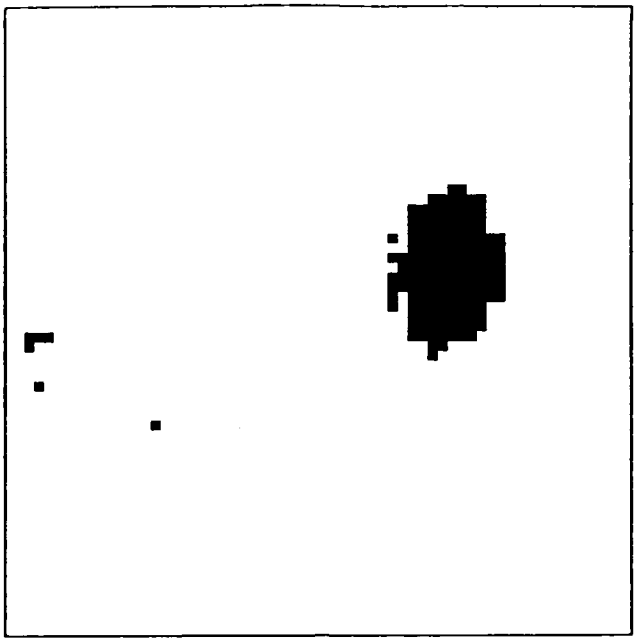


FIGURE 9b.



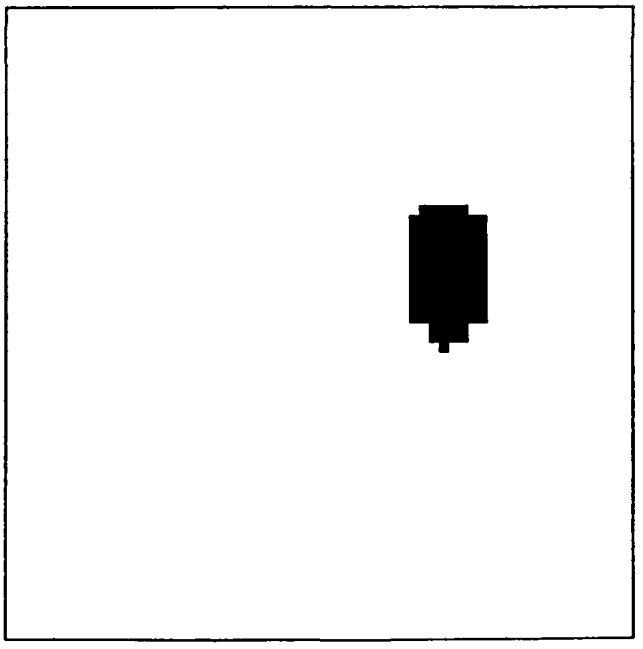


FIGURE 9c.

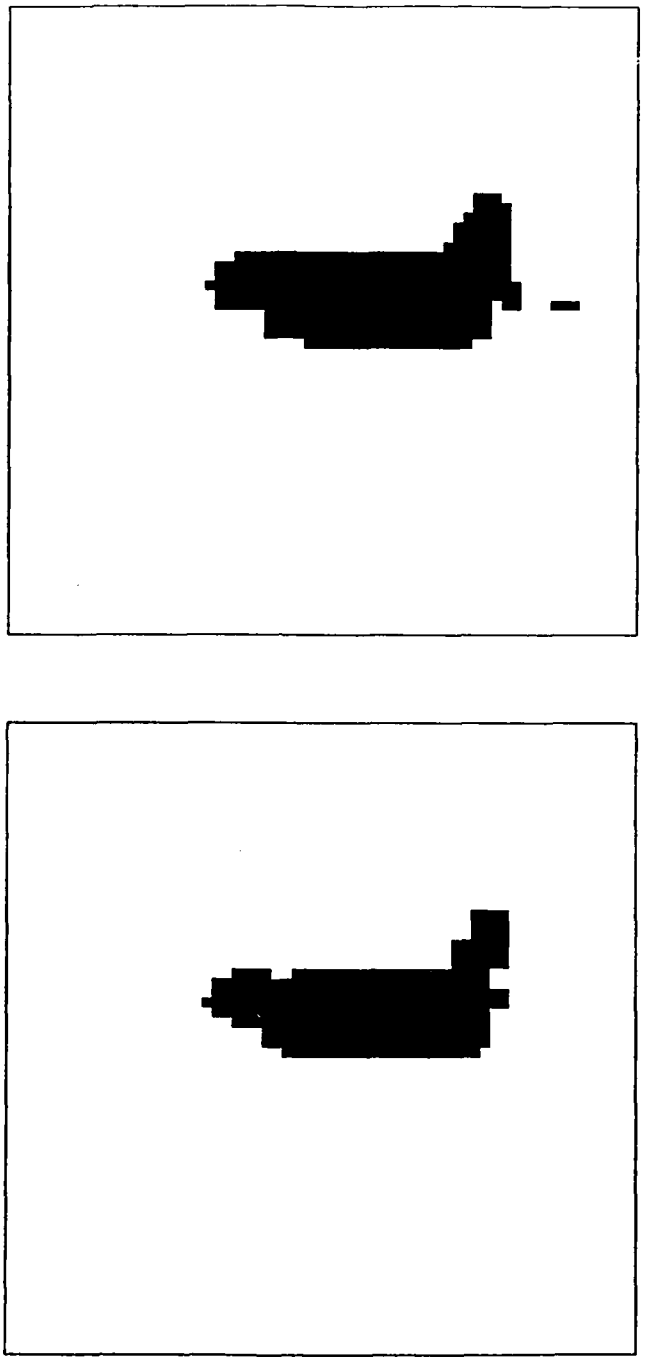
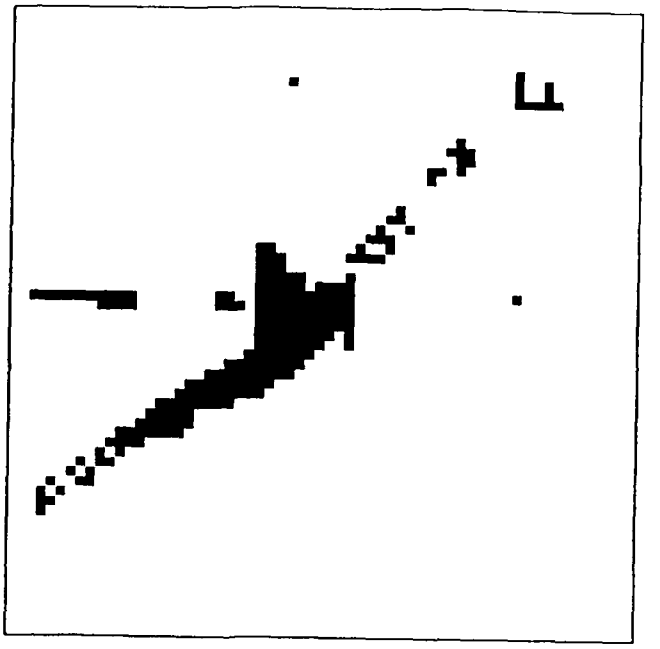


FIGURE 9d.



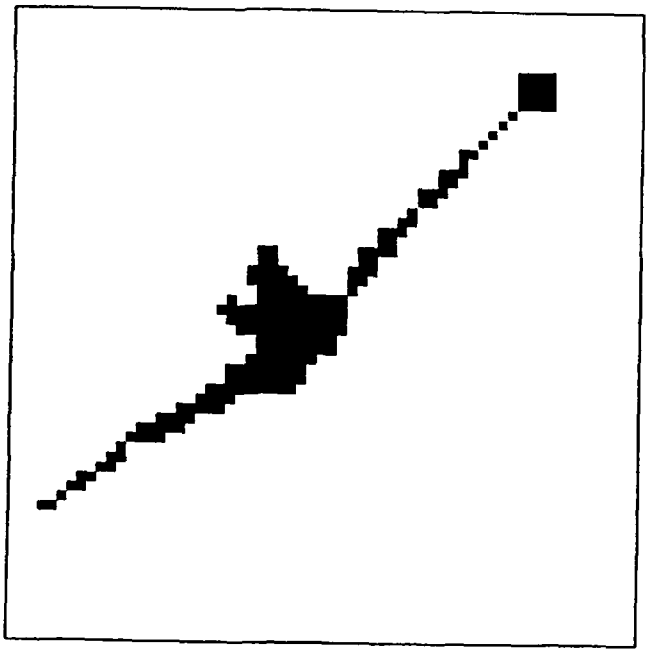


FIGURE 9e.

DISTRIBUTION

AUSTRALIA

Department of Defence

Defence Central
- Chief Defence Scientist
- AS, Science Corporate Management (shared copy)
- FAS Science Policy (shared copy)
- Director, Departmental Publications
- Counsellor, Defence Science, London (Doc Data Sheet Only)
- Counsellor, Defence Science, Washington (Doc Data Sheet Only)
- OIC TRS, Defence Central Library
- Document Exchange Centre, DISB (18 copies)

Aeronautical Research Laboratory
- Director
- Library
- Chief Aircraft Systems Division
- G. J. Saleeba

Army Office
- Scientific Adviser - Army
- Engineering Development Establishment
  - L. Dowling
- Director of Operational Requirements (DOR-A)
- Director of Artillery (DARTY)

SPARES (10 COPIES)

TOTAL (39 COPIES)

AL 149 DEPARTMENT OF DEFENCE

DOCUMENT CONTROL DATA

PAGE CLASSIFICATION
UNCLASSIFIED

PRIVACY MARKING

1a. AR NUMBER	1b. ESTABLISHMENT NUMBER	2. DOCUMENT DATE	3. TASK NUMBER
AR-005-522	ARL-SYS-TM-109	AUGUST 1988	ARM 86/066

4. TITLE
IMAGE SEGMENTATION FOR TARGET AIRCRAFT DETECTION

5. SECURITY CLASSIFICATION (PLACE APPROPRIATE CLASSIFICATION IN BOX(S) IE. SECRET (S), CONF. (C) RESTRICTED (R), UNCLASSIFIED (U)).

DOCUMENT	TITLE	ABSTRACT
U	U	U

6. NO. PAGES
26

7. NO. REFS.
7

8. AUTHOR(S)
J. Sandor and R.C. Creek

9. DOWNGRADING/DELIMITING INSTRUCTIONS
Originally released as Commerial In Confidence and has been downgraded and reprinted.

10. CORPORATE AUTHOR AND ADDRESS
AERONAUTICAL RESEARCH LABORATORY
P.O. BOX 4331, MELBOURNE VIC 3001

11. OFFICE/POSITION RESPONSIBLE FOR:
SPONSOR ARMY - OPERATIONAL REQ
SECURITY
DOWNGRADING
APPROVAL CSYD

12. SECONDARY DISTRIBUTION (OF THIS DOCUMENT)
Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH ASIDS, DEFENCE INFORMATION SERVICES BRANCH, DEPARTMENT OF DEFENCE, CAMPBELL PARK, ACT 2601

13a. THIS DOCUMENT MAY BE ANNOUNCED IN CATALOGUES AND AWARENESS SERVICES AVAILABLE TO ....
No limitations

13b. CITATION FOR OTHER PURPOSES (IE. CASUAL ANNOUNCEMENT) MAY BE [X] UNRESTRICTED OR [ ] AS FOR 13a.

14. DESCRIPTORS
Image processing
Target recognition
Missile simulators
Rapier anti aircraft system
Digital simulation
Anti aircraft missiles

15. DRDA SUBJECT CATEGORIES
0062F
0074A
0051F

16. ABSTRACT
A two stage method is presented for segmenting digitized images into aircraft and background. The approximate aircraft position is determined from the edge distribution and then each pixel within this region is classified by statistical comparison with the closest local region outside the search area.

PAGE CLASSIFICATION
UNCLASSIFIED
PRIVACY MARKING

THIS PAGE IS TO BE USED TO RECORD INFORMATION WHICH IS REQUIRED BY THE ESTABLISHMENT FOR ITS OWN USE BUT WHICH WILL NOT BE ADDED TO THE DISTS DATA UNLESS SPECIFICALLY REQUESTED.

16. ABSTRACT (CONT).

17. IMPRINT

AERONAUTICAL RESEARCH LABORATORY, MELBOURNE

18. DOCUMENT SERIES AND NUMBER	19. COST CODE	20. TYPE OR REPORT AND PERIOD COVERED
Aircraft Systems Technical Memorandum 109	71 6886	

21. COMPUTER PROGRAMS USED

22. ESTABLISHMENT FILE REF.(S)

23. ADDITIONAL INFORMATION (AS REQUIRED)